

APRENDIZAJE AUTOMÁTICO

ENTREGA 3

**TEMA: Modelo de Aprendizaje Automático
y Resultados**

**Integrante: CUSSI NICOLÁS
Profesor: MIRABETE MARTÍN**



**CENTRO POLITÉCNICO SUPERIOR
MALVINAS ARGENTINAS**



Entrega 3: Modelo de Aprendizaje Automático y Resultados

1. Origen de los Datos

En este proyecto se emplearon datos provenientes de fuentes oficiales y registros históricos sobre la producción de trucha arcoíris en Tierra del Fuego y zonas australes. Entre los conjuntos utilizados destacan: registros anuales de producción (MAGyP, FAO), datos climáticos (temperatura media anual del agua, aire y humedad), y registros de movimientos logísticos (alevinos y adultos). Para asegurar la integridad y representatividad, los datos fueron unificados, limpiados y en caso de ser necesario, completados mediante simulación, siguiendo las mejores prácticas de preparación de datos.

2. Análisis Exploratorio de Datos (EDA)

Se desarrolló un análisis exploratorio exhaustivo para comprender la dinámica y las tendencias de la producción. Entre los principales gráficos generados se incluyen:

- Evolución anual de la producción de trucha (línea de tiempo).
- Distribución de la producción en años considerados favorables vs. desfavorables (boxplot).
- Relación entre temperatura media del agua y producción (scatterplot).

Las estadísticas básicas (media, mediana, desvío estándar) y las visualizaciones permitieron detectar patrones de estacionalidad, correlaciones y la presencia de outliers.

3. Principales Conclusiones del EDA

El análisis evidenció que la producción presenta alta variabilidad interanual, estrechamente vinculada a las condiciones climáticas, especialmente a la temperatura del agua. Los años con temperaturas medias superiores a la mediana tienden a mostrar mejores niveles de producción. Asimismo, se observa que los movimientos logísticos y el manejo de los recursos piscícolas también tienen un impacto significativo.

4. Modelo de Aprendizaje Automático

Se optó por implementar modelos de Random Forest, tanto para regresión como para clasificación:

- **Variables empleadas:** producción histórica, temperatura media del agua, humedad, movimientos logísticos, y estacionalidad.
- **Ajuste de hiperparámetros:** el número de árboles y la profundidad máxima fueron optimizados mediante validación cruzada.
- **Entrenamiento y test:** se utilizó una división estratificada de los datos para asegurar la validez de las métricas obtenidas.

5. Métricas de Evaluación

El modelo fue evaluado usando:

- **Precisión, recall y F1-score** (para clasificación de condiciones favorables).
- **RMSE y MAE** (para predicción de la producción en toneladas).

El modelo alcanzó valores satisfactorios en estas métricas, confirmando su capacidad predictiva para el contexto regional y las variables empleadas.

6. Interpretación de Resultados y Conclusión

Los resultados obtenidos muestran que el modelo es útil para anticipar niveles de producción y distinguir entre años productivos y críticos. La integración de variables ambientales y logísticas resultó clave para mejorar el desempeño del modelo. Se concluye que la aplicación de Aprendizaje Automático aporta valor en la planificación y gestión de la acuicultura en Tierra del Fuego, permitiendo fundamentar decisiones estratégicas con base en datos objetivos y modelados.